George G. Vega Yon

Mobile	+1 (six two six) 381 8171
e-mail	g.vegayon@gmail.com
website	ggvy.cl
Code	github.com/gvegayon
Linkedin	www.linkedin.com/in/georgevegayon/

Education

- **Ph.D. in Biostatistics (with concentration in Stat. Comp.)** (2020)
  University of Southern California, USA
  Dissertation title:
  "Essays on Bioinformatics and Social Network Analysis: Statistical and Computational Methods for Complex Systems"

- **M.Sc. in Social Sciences (with concentration in Econometrics)** (2016)
  California Institute of Technology, USA

- **Master in Economics and Public Policy** (2011)
  Adolfo Ibáñez University, Chile

- **BS. in Business Administration** (with a minor in Political Science) (2010)
  Adolfo Ibáñez University, Chile

Professional Experience

- **University of Southern California, 2015–present** Department of Preventive Medicine *Research Programmer*. A senior staff member at USC's Department of Preventive Medicine, I work closely with both staff and students on various scientific projects. My responsibilities included: implement statistical methods using R/C++, analyze complex data using USC's High-Performance Computing cluster, conducting training sessions on statistical computing, and writing scientific papers.

- **Chilean Pension Supervisor, August 2011– August 2014** Research Division *Analyst*. Statistical and econometric analysis on the Chilean unemployment insurance, statistical software development, serving as a bridge between the IT and Research divisions.

- **Nodos Chile Social Network Analysis Ltda., January 2012–January 2014** *Partner*. Founding partner of one of the first applied SNA Consultancy Entrepreneurship in Chile.
- **Adolfo Ibáñez University, January 2011–June 2012.** School of Government *Adjunct Professor*. Taught Introductory courses of Economics, Microeconomics and Statistical computing with Stata.
- **Chilean Ministry of Social Planning, March 2011–December 2011.** Social Programs Monitoring *Analyst*. Survey and Analysis of the Government social programs supply and supporting the Monitoring Division with the Open-Data Initiative.

Peer Reviewed Publications

[1] Thomas W. Valente and **George G. Vega Yon**. "Diffusion/Contagion Processes on Social Networks". In: *Health Education & Behavior* 47.2 (2020). *contribution: I designed and implemented the data generating process used for the simulations of the paper*, pp. 235–248. URL: http://journals.sagepub.com/doi/10.1177/1090198120901497.

[2] **George G. Vega Yon**, Andrew Slaughter, and Kayla de la Haye. "Exponential random graph models for little networks". In: *Social Networks* (2020). *contribution: One of my main dissertation chapters, I developed both the statistical method and software, as well as conducted analyses and led the writing process*. Forthcoming.

[3] Kayla de la Haye, Heesung Shin, **George G. Vega Yon, et al.** "Smoking Diffusion through Networks of Diverse, Urban American Adolescents over the High School Period". In: *Journal of Health and Social Behavior* 60.3 (2019). *contribution: conducted part of the data analysis and created visualizations*, pp. 362–376. URL: https://doi.org/10.1177/0022146519870521.

[4] Thomas W. Valente, Heather Wipfli, and **George G. Vega Yon**. "Network influences on policy implementation: Evidence from a global health treaty". In: *Social Science and Medicine* 222 (2019). *contribution: Data collection, analysis, and visualization*, 188–197. URL: http://www.sciencedirect.com/science/article/pii/S0277953619300085.

[5] **George G. Vega Yon** and Brian Quistorff. "parallel: A command for parallel computing". In: *The Stata Journal: Promoting communications on statistics and Stata* 19.3 (Sept. 2019). *contribution: I first developed the Stata module, wrote the first draft, and led the writing process*, pp. 667–684. URL: http://journals.sagepub.com/doi/10.1177/1536867X19874242.

[6] **George G. Vega Yon** and Paul Marjoram. "fmcmc: A friendly MCMC framework". In: *Journal of Open Source Software* 4.39 (July 2019). *contribution: I implemented the algorithms and led the writing process*, p. 1427. URL: http://joss.theoj.org/papers/10.21105/joss.01427.

[7] **George G. Vega Yon** and Paul Marjoram. "slurmR: A lightweight wrapper for HPC with Slurm". In: *Journal of Open Source Software* 4.39 (July 2019). *contribution: I implemented the algorithms and led the writing process*, p. 1493. URL: https://joss.theoj.org/papers/10.21105/joss.01493.

[8] Brooke M. Bell, Donna Spruijt-Metz, **George G. Vega Yon, et al.** "Sensing eating mimicry among family members". In: *Translational Behavioral Medicine* 9.3 (May 2019). *contribution: I develop a novel statistical method for analyzing timestamped data using permutation tests and contributed to the writing process*, pp. 422–430.

[9] Jorge Fábrega Lacoa and **George G. Vega Yon**. "El impacto del rating televisivo sobre la actividad en Twitter: evidencia para Chile sobre la base del evento TELETÓN 2012". In: *Cuadernos.info* 33 (Dec. 2013). *contribution: Data Analysis and Visualization*, pp. 43–52. URL: http://cuadernos.info/index.php/CDI/article/view/533.

Work in Progress and Technical Reports

[1] **George G.. Vega Yon**, Duncan C. Thomas, John Morrison, et al. "On the automatic annotation of gene functions using observational data and phylogenetic trees". In: *bioRxiv* (May 2020). URL: https://www.biorxiv.org/content/early/2020/05/14/2020.05.14.095687.

[2] **George G. Vega Yon**. *Capital Necesario Unitario (CNU): Cálculo e Introducción Del Módulo de Stata CNU*. Working Papers 57. Superintendencia de Pensiones, Aug. 2014. URL: https://ideas.repec.org/p/sdp/sdpwps/57.html.

[3] Ximena Quintanilla, Isabel Poblete, **George G. Vega Yon, et al.** *Estudio Actuarial de los Fondos del Seguro de Cesantía*. Tech. rep. 2013.

[4] Andrea Repetto and **George G. Vega Yon**. *El Impacto de un Alza en la Cotización Previsional: Pensiones, Salarios y Empleo*. Tech. rep. 2013.

Software (selected)

[1] **George G. Vega Yon**. *rgexf: Build, Import and Export GEXF Graph Files* (2020). R package version

0.16.0. URL: https://CRAN.R-project.org/package=rgexf.
downloads 529K

[2] **George G. Vega Yon**, Thomas Valente. *netdiffuseR: Analysis of Diffusion and Contagion Processes on Networks* (2020). R package version 1.22.0. URL: https://github.com/USCCANA/netdiffuseR.
downloads 18K

[3] **George G. Vega Yon**, Kayla de la Haye. *ergmito: Exponential Random Graph Models for Small Networks* (2020). R package version 0.3-0. URL: https://cran.r-project.org/package=ergmito.
downloads 3126

[4] **George G. Vega Yon**. *slurmR: A Lightweight Wrapper for 'Slurm'* (2020). R package version 0.4-1. URL: https://CRAN.R-project.org/package=slurmR.
downloads 3194

[5] **George G. Vega Yon**. *fmcmc: A friendly MCMC framework* (2020). R package version 0.3-0. URL: https://CRAN.R-project.org/package=fmcmc.
downloads 4878

[6] **George G. Vega Yon**. *barry: your to-go motif accountant* (2020). C++ library version 0.0-1. URL: https://github.com/USCbiostats/barry.

[7] **George G. Vega Yon**. *pruner: Implementing the Felsenstein's Tree Pruning algorithm* (2020). C++ library version 0.0-1. URL: https://github.com/USCbiostats/pruner.

[8] **George G. Vega Yon**, Brian Quistorff. *parallel: Stata Module for Parallel Computing* (2019). Stata Module version 1.20.0. URL: https://github.com/gvegayon/parallel.

[9] **George G. Vega Yon**. *aphylo: Statistical Inference of Annotated Phylogenetic Trees* (2019). R package version 0.1.99. URL: https://github.com/USCbiostats/phylogenetic.

[10] **George G. Vega Yon**. *netplot: Beautiful graph drawing* (2019). R package version 0.0.9000. URL: https://github.com/USCCANA/netplot.

[11] **George G. Vega Yon**. *googlePublicData: Working with Google's 'Public Data Explorer' DSPL Metadata Files* (2017). R package version 0.16.1. URL: https://CRAN.R-project.org/package=googlePublicData.
downloads 24K

[12] **George G. Vega Yon**, Enyelbert Muñoz. *ABCoptim: Implementation of Artificial Bee Colony (ABC) Optimization* (2017). R package version 0.15.0. URL: https://CRAN.R-project.org/package=ABCoptim.
downloads 41K

[13] **George G. Vega Yon**. *twitterreport: Out-of-the-box analysis and reporting tools for twitter* (2016). R package version 0.16. URL: https://doi.org/10.5281/zenodo.44528.

Conferences

[1] **George Vega Yon**, Aileen Dinkjian, Sarah Hamm-Alvarez, et al. "ERGMito Statistical Models for Small Team Social Networks". In: SciTS 2020. (conference talk, slides/video). June 2020.

[2] **George Vega Yon**, Andrew Slaughter, and Kayla de la Haye. "Exact Statistics and Semi-Parametric Tests for Small Network Data". In: IC2S2, 2019. (conference talk, slides). July 2019.

[3] **George Vega Yon**, Andrew Slaughter, and Kayla de la Haye. "Exact Statistics and Semi-Parametric Tests for Small Network Data". In: Sunbelt 2019. (conference talk, slides). June 2019.

[4] **George Vega Yon** and Kayla de la Haye. "Small network statistics for the network science of teams". In: NetSciX 2019, SCL. (conference talk, slides). Jan. 2019.

[5] **George Vega Yon**. "Computacion de Alto Rendimiento con R". In: satRday Santiago 2018. (conference workshop, slides). Dec. 2018.

[6] **George Vega Yon** and Kayla de la Haye. "Big Problems for Small Networks: Small Network Statistics". In: NASN 2018, DC. (conference talk, slides). Nov. 2018.

[7] **George Vega Yon**, Kayla de la Haye, Hee-sung Shin, et al. "Diffusion of Smoking Initiation Among Diverse, Urban American Adolescents Over The High School Period". In: NASN 2017. (conference talk, slides). July 2017.

[8] **George Vega Yon** and Brian Quistorff. "Uncomplicated Parallel Computing with Stata". In: Stata Conference 2017. (conference talk, slides). July 2017.

[9] **George Vega Yon** and Thomas W. Valente. "Understanding Diffusion with netdiffuseR". In: NASN 2017. (conference workshop, slides). July 2017.

[10] **George Vega Yon** and Thomas W. Valente. "Network Diffusion of Innovations in R: Introducing netdiffuseR". In: IC2S2, 2016. (conference poster, slides). June 2016.

[11] **George Vega Yon** and Thomas W. Valente. "Network Diffusion of Innovations in R: Introducing netdiffuseR". In: useR! 2016. (conference talk, slides/video). June 2016.

[12] **George Vega Yon** and Thomas W. Valente. “Understanding Diffusion with netdiffuseR”. In: Sunbelt Conference, 2016. (conference workshop, slides/video). Mar. 2016.

[13] **George Vega Yon**. “Just tired of endless loops! or parallel: Stata module for parallel computing”. In: Stata Conference, 2013. (conference talk, slides). July 2013.

Invited Speaker

[1] “"Predicción de funciones genéticas utilizando evidencia experimental y árboles filogenéticos: Un modelo evolutivo" o "Ciencia de datos en la práctica"”. In: PUC Mathematical and Computational Engineering Seminars. (invited talk, slides). Apr. 2020.

[2] “A Brief Introduction to Using R for High-Performance Computing”. In: Orange County R Users Group. (invited talk, slides/video). Aug. 2019.

[3] “Big Problems for Small Networks: Statistical Analysis of Small Networks and Team Performance”. In: UCI Social Network Research Group. (invited talk, slides). Apr. 2019.

[4] “Big Problems for Small Networks: Statistical Analysis of Small Networks and Team Performance”. In: SONIC Speaker. (invited talk, slides/video). Mar. 2019.

[5] “Estadística de Redes y Econometría Espacial (con R)”. In: IMFD Summer School. (invited talk, slides/video). Dec. 2018.

Other Talks

[1] “Essays on Bioinformatics and Social Network Analysis: Statistical and Computational Methods for Complex Systems”. In: Doctoral Defense. (talk, slides). June 2020.

[2] “Essays on Bioinformatics and Social Network Analysis Statistical and Computational Methods for Complex Systems”. In: Biostats Seminars. (talk, slides/video). Jan. 2020.

[3] “slurmR workshop”. In: Happy Scientist Semminar Series. (workshop, slides). Jan. 2020.

[4] “Happy Scientist Seminar: Research Pipelines”. In: Happy Scientist Seminar. (talk, slides). Oct. 2019.

[5] “What drives social networks? A gentle introduction to exponential random graph models (with a focus on small networks)”. In: East LA R User Group. (talk, slides). June 2019.

[6] “Overview of Social Network Models”. In: NA. (talk, slides). Dec. 2018.

[7] "A brief introduction to using R for high-performance computing". In: East LA R User Group. (talk, slides). Nov. 2018.

[8] "R Bootcamp for Scientific Computing 2018". In: R Bootcamp for Scientific Computing. (workshop, slides). Aug. 2018.

[9] "Intro to R". In: USC's HPCC workshop. (workshop, slides). July 2018.

[10] "Introduction to R (for HPC users)". In: USC's HPCC workshop. (workshop, slides). July 2018.

[11] "Reproducible Research". In: HBR Students talk. (talk, slides). Sept. 2016.

Major Areas of Research Interest

Social Networks and Complex Systems
Scientific Software Development and High Performance Computing
Non-parametric Statistics
Statistical Methods Development

Honors and Services to the Profession

Awards

Travel Grant, Society of Young Network Scientist, 2019.
Fellowship, California Institute of Technology, 2014.
Honorable Mention (Posters Session) Chilean Economics Society, 2012.
Scholarship, Universidad Adolfo Ibáñes, 2006

Manuscript Review (Ad Hoc)

The Official Journal of The Society for Computational Economics
The R Journal
Social Networks
Journal of Mathematical Sociology
Journal of Open Source Software

Abstract Review

International Conference on Computational Social Science (2019)
International Conference on Computational Social Science (2020)
SUNBELT Conference (2016)

Book Review

"Microeconometrics and Matlab: An Introduction", by Adams, Clarke and Quinn, Oxford University Press, 2015.

"Mastering Gephi Network Visualization", by Ken Cherven, Packt Publishing, 2015.

"Network Graph Analysis and Visualization with Gephi", by Ken Cherven, Packt Publishing, 2013.

Misc

Founder of the (first) R Users Group in Chile (2013)

Co-organizer of the East LA R User Group (LAERUG).

last update: July 22, 2020

https://ggvy.cl